A METHOD OF RESTORING PARTIALS OF A SOUND SIGNAL

The present invention relates to the field of telecommunications and in particular to the field of digitally processing a sound signal and to the harmonic

5   representation of a sound signal.

In harmonic modeling of digital audio signals, the sound signal is represented by a set of oscillators whose parameters (frequency, amplitude, phase) vary slowly over time. The harmonic analysis comprises short-term

10   time/frequency analysis for determining the values of these parameters followed by extraction of peaks and then tracking of partials.

The signal to be modeled is divided into frames of ℓ samples (typically $\ell=1024$). A short-term time/frequency

15   analysis module (which typically executes a Fourier transform) calculates the short-term spectrum of the signal for each frame. A module for extracting peaks retains only the peaks that are the most pertinent *a priori*, one criterion being keeping only the highest

20   energy peaks, for example. A third and final module attempts to link the peaks with each other over time, i.e. from one frame to another, to form the partials. During its life cycle each partial corresponds to one oscillator.

25   That type of analysis and representation may be used in particular during bit rate reduction coding, parametric coding (which processes three aspects of the signal: transients, sinusoids, noise), separation and indexing of sound sources, and restoration of sound

30   files.

It is accepted at present that the best quality is achieved when synthesizing partials by using phase interpolation techniques proposed by Robert J. McAulay and Thomas F. Quatieri in the paper "Speech

35   Analysis/Synthesis Based on a Sinusoidal Representation", IEEE Transactions on Acoustics, Speech and Signal Processing, pp. 744-754, 1986, or by Laurent Girin,

Sylvain Marchand, Joseph di Martino, Axel Röbel and Geoffroy Peeters in the paper "Comparing the order of a Polynomial Phase Model for the Synthesis of Quasi-Harmonic Audio Signals", WASPAA, New Paltz, NY, USA, October 2003. Those techniques are used to synthesize a partial from a peak (A_i, f_i, φ_i) to a peak $(A_{i+1}, f_{i+1}, \varphi_{i+1})$ by calculating all the intermediate phases using third or fifth order polynomials, the frequencies being deduced by derivation. Third order interpolation is used when only the start and end frequencies and phases are known. Fifth order interpolation is used when the second order variations of the phase are also known (these are equivalent to first order variations of frequency since by definition frequency is the derivative of phase).

Synthesizing a partial between the peaks $P_i(A_i, f_i, \varphi_i)$ and $P_{i+1}(A_{i+1}, f_{i+1}, \varphi_{i+1})$ consists in calculating the values $p(n)$ of the partial between the frames i and i+1:

$$p_i(n) = p(li+n) = A_i(n)\cos(\varphi_i(n)), \quad n = 0,...,\ell-1 \tag{1}$$

To this end, it is known in the art to calculate all the intermediate phases using one of the following two interpolation methods.

For third order interpolation according to McAulay, the phase is calculated from the following expression, in which Te is the sampling period:

$$\varphi_i(n) = \varphi_i + 2\pi f_i nTe + \alpha(nTe)^2 + \beta(nTe)^3 \tag{2}$$

The two unknowns α and β are calculated by solving a system of equations in $(f_i, \varphi_i, f_{i+1}, \varphi_{i+1})$. The frequencies are deduced by differentiation:

$$2\pi f_i(n) = 2\pi f_i + 2\alpha nTe + 3\beta(nTe)^2 \tag{3}$$

For fifth order interpolation according to Girin *et al.*, the first order variations δf_i and δf_{i+1} of the frequency at the peaks P_i and P_{i+1} are assumed to be known. The phase is then calculated from the following expression:

$$\varphi_i(n) = \varphi_i + 2\pi f_i nTe + \frac{\delta f_i}{2}(nTe)^2 + \beta(nTe)^3 + \gamma(nTe)^4 + \delta(nTe)^5 \tag{4}$$

The three unknowns β, δ, γ are calculated by solving a system of equations in $(f_i, f_{i+1}, \varphi_i, \varphi_{i+1}, \delta f_i, \delta f_{i+1})$. The frequencies are deduced by differentiation:

$$2\pi f_i(n) = 2\pi f_i + \delta f_i nTe + 3\beta(nTe)^2 + 4\gamma(nTe)^3 + 5\delta(nTe)^4 \qquad (5)$$

5   For various reasons, it may happen that certain partials in the signal are absent, corrupted, or discontinuous at the end of analysis and/or at the beginning of synthesis. For example, they may be absent at the input of the decoder in an Internet sound program
10   broadcast application in the event of loss of packets, they may be corrupted if the signal to be analyzed is suffering interference from an unwanted signal (noise, click, other signal, etc.), or they may be discontinuous if their energy is too low for them to be correctly
15   detected continuously. To create a synthesized signal as close as possible to the original signal it is then necessary to restore the missing peaks. This necessitates creating peaks each characterized by an amplitude, a frequency, and a phase.
20   The above prior art interpolation techniques are used to synthesize the portions corresponding to the missing peaks and to restore the partials.

However, those prior art interpolation techniques are adapted to use in the short-term, i.e. over a period
25   of less than 10 milliseconds (ms). For longer periods, the re-synthesized signal is often very different from the original signal and disagreeable artifacts may appear. Those techniques ensure phase continuity between the existing peaks and the restored peaks but are not
30   able to control the induced frequencies resulting from equations (3) and (5). That effect increases in direct proportion to the interpolation distance.

One object of the invention is to propose an alternative solution to the problem of restoring a
35   missing portion identified as that of a partial, in particular if the missing portion corresponds to a long

period (greater than 10 ms), for which the prior art techniques are relatively ineffective.

Accordingly, the technical problem to be solved by the present invention is that of proposing a method of restoring missing portions of partials of a sound signal during harmonic analysis in which the sound signal is divided into time frames to which time/frequency analysis is applied that supplies successive short-term spectra represented by sample frequency frames, the analysis further consisting in extracting spectrum peaks in the frequency frames and linking them together over time to form partials, this method being an alternative to the prior art solutions.

In accordance with the present invention, one solution to the stated technical problem consists in that said method of restoring a partial between a peak P_i and a peak P_{i+N} whose frequency ω and phase are known is characterized in that it comprises the steps of:

· estimating the frequency $\hat{\omega}$ of each of the missing peaks P_{i+1} to P_{i+N-1} of this partial;

· calculating the phase $\hat{\varphi}$ from peak to peak, from the phase of the peak P_i to that of the peak P_{i+N}, for all the frequencies $\hat{\omega}$ previously estimated;

· calculating the phase error $err\varphi$ between the calculated phase $\hat{\varphi}$ and the known phase at the same peak P_{i+N};

· correcting each calculated phase $\hat{\varphi}$ by a value that is a function of the phase error $err\varphi$.

A method of the invention differs from the prior art methods in that it offers finer control of the frequency of the missing peaks and subsequent calculation of the corresponding phases to ensure continuity with the phases of the existing peaks. Accordingly, a method of the invention re-synthesizes signals corresponding to the missing partial portions without artifacts, in contrast to the prior art methods described above.

A method of the invention also has the advantage of reconstructing a signal that is closer, in terms of the reconstruction error, to the original signal than is the signal obtained by the prior art methods.

5   Finally, a method of the invention has the advantage of using an algorithm of low complexity.

The invention further consists in a synthesizer for synthesizing a sound signal for implementing a method of restoring a partial between a peak P_i and a peak P_{i+N}, for

10   example an audio decoder or a parametric coder adapted to use a method of the invention.

The invention further consists in a computer program product loadable directly into the internal memory of the above synthesizer or group of synthesizers and comprising

15   software code portions for executing steps of a method according to the invention when the program is executed on the synthesizer or group of synthesizers.

The invention further consists in a medium usable in the above synthesizer or group of synthesizers on which

20   there is stored a computer program product loadable directly into the internal memory of the synthesizer or group of synthesizers and comprising software code portions for executing steps of a method according to the invention when the program is executed on the synthesizer

25   or group of synthesizers.

Other features and advantages of the invention become apparent in the course of the following description, which is given with reference to the appended drawing, which is provided by way of

30   non-limiting example.

Figure 1 is a flowchart of one example of the invention.

Figure 2 is a diagram of one example of the use of a method of the invention.

35   A method 1 of the invention proceeds in the following manner, described here with reference to the Figure 1 flowchart. The method consists in restoring a

partial between a peak P_i and a peak P_{i+N} whose frequencies ω and phases φ are known.

In a first step 2, the method estimates the frequency $\hat{\omega}$ and the amplitude A of each of the missing peaks P_{i+1} to P_{i+N-1}, for example by linear prediction or interpolation methods known in the art.

Consider a partial consisting of a succession of linked peaks $P_i(A_i, \omega_i, \varphi_i)$ known at times iT and characterized by:

A_i, the amplitude of the peak at the time iT;

ω_i, the frequency of the peak at the time iT; and

φ_i, the phase of the peak at the time iT, modulo 2π.

The frequency of the missing peaks between the peaks P_i and P_{i+N} is estimated by means of linear interpolation between ω_i and ω_{i+N}, for example, or linear past or future prediction, as described in the paper "Enhanced Partial Tracking using linear Prediction", Mathieu Lagrange, Sylvain Marchand, Martin Raspaud and Jean-Bernard Rault, Proceedings of the Digital Audio Effects (DAFx) Conference, pp141-146, Queen Mary College, University of London, UK, September 2003, for example, or by means of a weighted past or future combination.

The amplitude A of the missing peaks is estimated by linear interpolation between A_i and A_{i+N}, for example, linear past or future prediction, or weighted past or future combination.

In a second step 3, the method calculates the phase $\hat{\varphi}$ from peak to peak, from the phase of the peak P_i to that of the peak P_{i+N}. This calculation is effected for each of the frequencies ω previously estimated.

Let φ_i and ω_i be the starting phase and frequency and $\{\hat{\omega}_{i+1}, ..., \hat{\omega}_{i+N-1}\}$ estimated frequencies in the range to be reconstructed. To extend the partial between the peak P_i and the peak P_{i+N} the phase is calculated from the following expression:

$$\hat{\varphi}_{i+n} = \mathrm{mod}\left(\varphi_i + \sum_{j=1}^{n} \frac{\hat{\omega}_{i+j} + \hat{\omega}_{i+j-1}}{2}T, 2\pi\right), \quad n = 1,...,N \qquad (6)$$

To avoid generating discontinuities that would compromise the quality of the re-synthesis, it is necessary to obtain at the time i+N a reconstructed phase $\hat{\varphi}_{i+N}$ equal to φ_{i+N}. The data in the above expression (6) being either approximate or predicted, it is statistically impossible to obtain this equality. Consequently, the subsequent steps of the method divide the phase error *errφ* calculated at the time i+N between all the missing peaks P_{i+1} to P_{i+N-1} previously reconstructed.

In a third step 4, the method calculates the phase error *errφ* between the calculated phase $\hat{\varphi}_{i+N}$ and the known phase φ_{i+N} at the same peak P_{i+N}. This calculation may use the following system of equations:

$$\text{if } |\varphi_{i+N} - \hat{\varphi}_{i+N} + 2\pi| < |\varphi_{i+N} - \hat{\varphi}_{i+N}|, \quad err\varphi = \varphi_{i+N} - \hat{\varphi}_{i+N} + 2\pi \qquad (7)$$

$$\text{if } |\varphi_{i+N} - \hat{\varphi}_{i+N} - 2\pi| < |\varphi_{i+N} - \hat{\varphi}_{i+N}|, \quad err\varphi = \varphi_{i+N} - \hat{\varphi}_{i+N} - 2\pi \qquad (8)$$

$$\text{else } err\varphi = \varphi_{i+N} - \hat{\varphi}_{i+N} \qquad (9)$$

In a fourth step 5, the method corrects each calculated phase $\hat{\varphi}_{i+n}$ by a value that is a function of the phase error *errφ*. The phase error calculated at the time i+N is typically divided uniformly between the calculated phases in accordance with the following expression:

$$\hat{\varphi}corrected_{i+n} = \mathrm{mod}\left(\hat{\varphi}_{i+n} + err\varphi\frac{n}{N}, 2\pi\right) \quad n = 1,...,N-1 \qquad (10)$$

The distribution need not be uniform, and may conform to a non-linear law, for example.

The Figure 2 example of use consists in restoring partials by means of the method 1 of the invention at the time of harmonic analysis of a sound signal, for example during parametric coding. The sound signal s(n) is represented by a set of oscillators whose parameters (frequency, amplitude) vary slowly over time. In the conventional way, the harmonic analysis includes short-term time/frequency analysis 6 for determining the

values of these parameters, followed by extraction of peaks 7 followed by tracking 8 of partials. Detection 9 of gaps in the partials precedes restoring partials by the method 1 of the invention. The peaks

5    $P_{i+n}\left(\hat{A}_{i+n}, \hat{\omega}_{i+n}, \hat{\varphi}_{i+n}\right)$ reconstructed by executing the method 1 are

then treated as peaks resulting from the harmonic analysis and additive synthesis 10 of the signal corresponding to the partial restored from these reconstructed peaks may be effected by one of the prior

10   art (third or fifth order) phase interpolation methods, for example.